Splunk N’ Box
Splunk Multi-Site Clusters In 20 Minutes or Less!

Mohamad Hassan | Sales Engineer

9/25/2017 | Washington, DC
Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Listen to Your Data, The Engine for Machine Data, Splunk Cloud, Splunk Light and SPL are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2017 Splunk Inc. All rights reserved.
About Me…

- Splunk SE (2 years), St. Louis, MO
- 23+ Years IT experience
- Splunk Admin/5TB (5 years)
- C/Unix Developer (4 years)
- Unix Admin
- Security Architect/Incident Responder
- Large Scale Deployments
- Creator of the first Email-To-Pager gateway (ePage) 1998
  - https://www.linkedin.com/in/hassanmohamad/
  - https://www.splunk.com/blog/2016/05/05/high-performance-syslogging-for-splunk-using-syslog-ng-part-1.htm
What Are We Solving?

- I don’t have the time to build a test environment
- I don’t have the budget (most testing done on my laptop)
- I just want to focus on Splunk and don’t want to learn docker/VM
- I need a training lab to teach Splunk
- I share the lab with other teams
- Cannot “truly” replicate my production environment in my lab
Splunk N’ Box Journey

- One year in the making
- 4500+ lines of bash
- 98 functions
- Started as 20 lines
- Optimize for MacOS
- User-feedback driven features
Imagine What You Can Build In 40 Minutes!
Splunk n’ Box Possible Use Cases

- Classroom or Search Parties
- Fully replicate production environment in your lab
- Learn Clustering without learning docker
- Test upgrades, new features or configurations
- Test integration with 3rd party (MySQL, Hadoop, etc.)
- Test apps in distributed environments
- Offline splunk demos (internal use)
- Splunk certification
Docker Quick Overview

- Began as an open-source implementation of the deployment engine which powers dot Cloud
- A platform for managing Linux Containers
- Rich set of API
- Small footprint and fast
- Very active user community
- Easy to script
- Fully Automated, Easy To Deploy, Quickly Scale
- Hosts provisioning: days -> minutes
Linux Kernel Features *used* by Docker

- **Namespaces**
  - (mnt, pid, net, ipc, uts/hostname, user ids)

- **cgroups**
  - (cpu, memory, disk, i/o - resource management)

- **AppArmor, SELinux**
  - (security/access control)

- **seccomp**
  - (computation isolation)

- **chroot**
  - (file system isolation)
VM vs. Container

Virtual Machine

Hypervisor
Host OS
Physical OS

Container

Docker Engine
Host OS
Physical OS
Where Is Docker In This Spectrum?

Splunk in VM  Splunk in Docker  Splunk Native
Docker Architecture

- CLI
  - docker build
  - docker pull
  - docker run

- Docker daemon
  - DOCKER_HOST
  - Containers
  - Images

- Images Repo
  - Registry
    - Docker Images
    - NGINX
Docker Tools

- **Docker Compose**: create and manage multi-container architectures
- **Kitematic**: Simple application for managing Docker containers on Mac and Windows
- **Docker Swarm**: orchestrating tool to provision and schedule containers
- **Docker Machine**: provision hosts and install Docker on them
- **VBox/Xhyve/Hyper-V**: Virtualization software to run Docker host for Mac and Windows
Dockerfile?

- **Dockerfile is instructions to build Docker image**
  - How to run commands
  - Add files or directories
  - Create environment variables
  - What process to run when launching container

- Result from building Dockerfile is Docker image
- Use Docker image to create container(s)
**Splunk N’ Box Features**

1. Extensive error checking during startup & while building containers
2. Adaptive load control during cluster builds
3. Built-in dynamic hostnames and IPs allocation (DHCP like)
4. IP aliases binding. No need to translate Splunk ports or proxy (nginx)
5. Automatically create & configure large number of Splunk hosts very fast
6. Different levels of logging (show docker commands executed)
7. Fully configured multi & single site cluster builds (LM, CM, DEP)
8. Optimize for performance
9. Menu driven & automatic code upgrade
10. Splunk DEMOs automation (no Docker knowledge required)
11. Linux, MacOS, Windows WSL (Ubuntu Linux subsystem), AMZ EC2
12. Custom login screen (Lab & Search Parties scenario)
Configuring The Script

ETH_OSX="lo0" #default interface to use with OSX
ETH_LINUX="eno1" #default interface to use with Linux
GREP_OSX="/usr/local/bin/ggrep"
GREP_LINUX="/bin/grep"

START_ALIAS_LINUX="192.168.1.100";
END_ALIAS_LINUX="192.168.1.254"
START_ALIAS_OSX="10.0.0.100";
END_ALIAS_OSX="10.0.0.254"

DNSSERVER="192.168.1.100"
LIC_FILES_DIR="licenses_files";
VOL_DIR="docker-volumes"

SPLUNK_IMAGE="splunkbbox/splunk_6.6.2"
RFACTOR="3";
SFACCTOR="2"

STD_IDXC_COUNT="3" #default IDXC count
STD_SHC_COUNT="3" #default SHC count
DEP_SHC_COUNT="1"

DEFAULT_SITES_NAMES="STL LON HKG"
Host (Container) Naming Rules

IDX : Indexer
SH  : Search Head
DS  : Deployment Server
LM  : License Master
CM  : Cluster Master
DEP : Search Head Cluster Deployer
HF  : Heavy Forwarder
UF  : Universal Forwarder
DMC : Distributed Management Console (splunk 6.5 name changed to Monitoring Console)
MacOS Notes:

1. **Default docker settings on MacOS are limited**
   Please change to take advantage of all available memory and CPU (under preferences).

2. **Performance on MacOS is noticeably less than Linux**
   So be aware that you may not be able to bring up as many containers with similar hardware resources.

3. **Hosts will not be reachable from outside your laptop**
   Containers will bind to local loopback interface IP aliases on docker-host (i.e., your laptop). This is not the case in Linux runs.

4. **Do not run any local splunkd instances on the docker-host**
   It will prevent Docker containers from starting due to network interface binding conflict.

5. **Do not use older boot2docker stuff**
   If you Google OSX Docker install, you will see references to Oracle VirtualBox and boot2docker everywhere. Starting with Docker 1.12 Oracle VBOX is replaced with small new hypervisor called xhyve. Boot2docker is replaced with Moby (tiny Linux).
My LAB

STAND ALONE LAB (25 containers)

- MacBook Pro
  - MacOS Sierra 10.12.2
  - 16G RAM
  - Intel Core i7

SEACH PARTY/CLASSROOM (80 containers)

- Netgear R800 router
  - SSID: splunk_n_box

- MacOS Sierra 10.12.2
  - 16G RAM
  - Intel Core i7

- Ubuntu 16.0.4

- Netgear R800 router
  - SSID: splunk_n_box

- 32G RAM, 1TB SSD
  - Intel Core i7

- Ubuntu 16.0.4
Screen Shots
Startup Checks
Main Menu

1) Manage All Containers & Images
2) Manage Lunch & Learn Containers
3) Manage Splunk Clusters
4) Manage Splunk Demos [**internal use only**]
5) Manage 3Rd Party Containers & Images [**under construction**]
6) Manage System
7) Change Log Level
8) Help
9) Quit

Enter your choice [1-6]
Manage All Containers & Images

Splunk n' Box v4.2.2.9: MAIN MENU -> SPLUNK MENU

- DOCKER: [ver:17.03.1-ce cpu:8 mem:15GB] OS: [FreeMem:5.2GB Load:3.37] Image: [splunkbox/splunk_6.5.3] LogLevel:[3]

Manage Images:
1) SHOW all images details [docker rmi --force $(docker images)]
2) REMOVE image(s) to recover disk-space (will extend build times) [docker rmi --force $(docker images)]
3) DEFAULT Splunk images [currently: splunkbox/splunk_6.5.3]

Manage Containers:
C) CREATE generic Splunk container(s) [docker run ...]
L) LIST all containers [custom view]
P) STOP container(s) [docker stop $(docker ps -aq)]
T) START container(s) [docker start $(docker ps -a --format "{{Names}}")]
D) DELETE container(s) & Volumes(s) [docker rm -vf $(docker ps -aq)]
H) HOSTS grouped by role [works only if you followed the host naming rules]

Manage Splunk:
E) RESET all splunk passwords [changeme --> hello] [splunkd must be running]
N) LICENSES reset [copy license file to all instances]
U) SPLUNK instance(s) restart

Manage system:
B) BACK to MAIN menu
H) HELP!

Enter choice (? for help):
Manage Splunk Clusters
<table>
<thead>
<tr>
<th>Host(container)</th>
<th>State</th>
<th>Splunkd Ver</th>
<th>Internal IP</th>
<th>Image used</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) MONITOR</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.8.2</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-205-179-124.compute-1.amazonaws.com:8088">http://ec2-34-205-179-124.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>2) SITE01CM01</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.5</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-196-175-111.compute-1.amazonaws.com:8088">http://ec2-34-196-175-111.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>3) SITE01DEP01</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.10</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-200-192-75.compute-1.amazonaws.com:8088">http://ec2-34-200-192-75.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>4) SITE01DMC01</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.3</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-195-168-186.compute-1.amazonaws.com:8088">http://ec2-34-195-168-186.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>5) SITE01HF01</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.6</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-197-121-37.compute-1.amazonaws.com:8088">http://ec2-34-197-121-37.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>6) SITE01IX01</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.7</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-197-174-106.compute-1.amazonaws.com:8088">http://ec2-34-197-174-106.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>7) SITE01IX02</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.8</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-198-136-6.compute-1.amazonaws.com:8088">http://ec2-34-198-136-6.compute-1.amazonaws.com:8088</a></td>
</tr>
<tr>
<td>8) SITE01IX03</td>
<td>Up</td>
<td>Running 6.5.3</td>
<td>172.18.0.9</td>
<td>splunk_6.5.3</td>
<td><a href="http://ec2-34-198-159-78.compute-1.amazonaws.com:8088">http://ec2-34-198-159-78.compute-1.amazonaws.com:8088</a></td>
</tr>
</tbody>
</table>
Finished Run

Creating hosts

[SITE01-DEP01:10.0.0.108] Creating new splunk docker container OK!
[SITE01-SH01:10.0.0.109] Creating new splunk docker container OK!
[SITE01-SH02:10.0.0.110] Creating new splunk docker container OK!
[SITE01-SH03:10.0.0.111] Creating new splunk docker container OK!

Finished creating hosts

Starting PHASE2: Converting generic SH hosts into SHC

[SITE01-DEP01] Configuring Deployer ...
[SITE01-SH01] Making cluster member...
[SITE01-SH02] Making cluster member...
[SITE01-SH03] Making cluster member...
[SITE01-SH03] Configuring as Captain (last SH created)...
[SITE01-SH03] ==> Checking SHC status (on captain)... OK!

Execution time for create_single_shc(): [2:51]
Number of Splunk config commands issued: [51]

Number of Splunk Commands Used to Build The Cluster = 74
Total execution time for build_single_site = **10:19** minutes

Hit <ENTER> to continue...
Real Performance Numbers

Intel NUC Skull 32G/SSD/Intel i7/Ubuntu 16.04:

- **Basic Splunk container:** (custom web.conf, pass changed, license file)
  - 2 splunk commands  20 seconds

- **1-Site cluster:** each site (3-IDX, 3-SH, 3-DEP), 1-CM, 1-LM
  - 224 splunk commands  18:10 minutes

- **4-Site cluster:** each site (10-IDX, 5-SH, 4-DEP), 1-CM, 1-LM
  - 625 splunk commands  38:58 minutes
FAQ

1. Can you run different Splunk version?
2. Do I need valid Splunk licenses?
3. Where is vi, ifconfig, sshd? How do I login into the container (docker-ssh)?
4. Can I run this script in production?
5. Is this script supported by Splunk?
6. Does it run on other Linux distribution beside Ubuntu or OSX?
7. Does it run on Windows?
8. Is the script using docker swarm?
9. Why there is a hypervisor used with MacOS/Windows?
10. Can I run this script inside a VM?
Learn docker in 10 minutes video:
https://www.youtube.com/watch?v=YFl2mCHdv24

Full & detailed Splunk n’ Box video (google splunk n box):
https://www.youtube.com/watch?v=k1WmnIWa4lo&feature=youtu.be

Ant Lefebvre/Presidio Splunk n’ Box on USB stick
alefebvre@presidio.com

https://youtu.be/qTAS1gvIGxM
Online Demo Video

https://www.youtube.com/watch?v=q1mRrpX-iLE
1. No need to learn docker/VM. Focus on learning Splunk
2. Quick and easy way to learn clustering
3. Potentially a game changer
4. 10-15 mins installation time
5. Keep the feedback coming
Thank You

Don't forget to rate this session in the .conf2017 mobile app